ARI Research Note 91-17

The Usefulness of Project A Spatial Tests for Predicting Comprehensive Performance Measures

Henry H. Busciglio
U.S. Army Research Institute

Selection and Classification Technical Area
Michael G. Rumsey, Chief

Manpower and Personnel Research Laboratory

Zita M. Simutis, Director

January 1991

United States Army
Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

U.S. ARMY RESEARCH INSTITUTE FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

JON W. BLADES
COL, IN
Commanding

Technical review by

Mark Curtis Young

NOTICES

DISTRIBUTION: This report has been cleared for release to the Defense Technical Information Center (DTIC) to comply with regulatory requirements. It has been given no primary distribution other than to DTIC and will be available only through DTIC or the National Technical Information Service (NTIS).

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The views, opinions, and findings in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other authorized documents.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS: --
2a. SECURITY CLASSIFICATION AUTHORITY: --	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: --	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): ARI Research Note 91-17	5. MONITORING ORGANIZATION REPORT NUMBER(S): --
6a. NAME OF PERFORMING ORGANIZATION: U.S. Army Research Institute; 6b. OFFICE SYMBOL (If applicable): PERI-RS	7a. NAME OF MONITORING ORGANIZATION: --
6c. ADDRESS (City, State, and ZIP Code): 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	7b. ADDRESS (City, State, and ZIP Code): --
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: U.S. Army Research Institute for the Behavioral and Social Sciences; 8b. OFFICE SYMBOL (If applicable): PERI-R	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: --
8c. ADDRESS (City, State, and ZIP Code): 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 63007A; PROJECT NO. 792; TASK NO. 2207; WORK UNIT ACCESSION NO. H1

11. TITLE (Include Security Classification)
The Usefulness of Project A Spatial Tests for Predicting Comprehensive Performance Measures

12. PERSONAL AUTHOR(S)
Busciglio, Henry H.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 90/10 TO 90/12	1991, January	15

16. SUPPLEMENTARY NOTATION
--

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Project A spatial tests, Predicting comprehensive performance measures, Regression analysis

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The Army's Project A developed and validated measures of abilities other than the general cognitive domain covered by the Armed Services Vocational Aptitude Battery (ASVAB). Busciglio (1990) analyzed data from the 1985 Concurrent Validation phase of Project A and found that Project A spatial scores substantially incremented the predictive validity of ASVAB. However, a number of questions were left unanswered. How useful are spatial abilities to performance in various entry-level Army jobs? Do spatial tests remain valid predictors after the ASVAB subtests have been added to prediction equations? Busciglio's (1990) original sample of 4,039 individuals in nine entry-level MOS was again used. Predictors were the ASVAB and the Project A spatial tests. Criteria were comprehensive measures of job performance and included written and hands-on tests of common and MOS-specific tasks. Backward stepwise multiple regression analyses were performed to answer the research questions. In comparing the results with those of Busciglio (1990), it was found that the spatial tests accounted for approximately as much criterion variance as did the ASVAB tests,

(Continued)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Michael G. Rumsey	22b. TELEPHONE (Include Area Code): (703) 274-8275; 22c. OFFICE SYMBOL: PERI-RS

DD Form 1473, JUN 86 Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

ARI Research Note 91-17

19. ABSTRACT (Continued)

indicating the importance of spatial skills to the performance of many Army jobs. Also, several of the spatial tests were especially strong predictors of the criterion measures across MOS, suggesting that they would lead to improved prediction of job performance if incorporated into the Army's selection composites. A number of methodological considerations are noted as an aid to interpreting these results.

THE USEFULNESS OF PROJECT A SPATIAL TESTS FOR PREDICTING COMPREHENSIVE PERFORMANCE MEASURES

CONTENTS

LIST OF TABLES

The Usefulness of Project A Spatial Tests For Predicting Comprehensive Performance Measures

Introduction

The Army's Project A is a long-term, comprehensive effort to improve the selection and classification of enlisted personnel. One objective of this effort was to validate the Armed Services Vocational Aptitude Battery (ASVAB, 1984), the traditional instrument for assessing the general cognitive skills of potential soldiers. Previous analyses of Project A data (Campbell, 1988) demonstrated that the ASVAB is useful for predicting first-tour performance. Therefore, the ASVAB serves as a baseline against which the marginal utility of other tests for selection and classification is judged.

Another objective of Project A was to develop and validate measures of abilities other than the general cognitive domain covered by ASVAB. For example, Project A staff members hypothesized that measures of spatial and perceptual-psychomotor abilities would account for criterion variance which was not predictable from ASVAB scores. In pursuit of this objective, the new Project A predictors and criterion measures were administered to approximately 9,500 soldiers in 19 entry-level Military Occupational Specialties (MOS) in the 1985 Concurrent Validation phase of Project A.

In a previous analysis of the 1985 Concurrent Validity data, Busciglio (1990) found that Project A spatial and perceptual-psychomotor test scores substantially improved the prediction of many criteria, most notably three comprehensive measures of success: total score on written tests of school and job knowledge, General Soldiering Proficiency, and Core (i.e., MOS-specific) Technical Proficiency. The statistical procedure used was a series of backward stepwise regressions in which the ASVAB subtests were entered into the equation in a block and nonsignificant subtests were removed and the Project A tests were entered to determine the extent to which they were significant as incremental predictors.

Busciglio's (1990) results left a number of important research questions unanswered. The first of these is basically theoretical: How useful are spatial abilities to performance in various entry-level Army jobs? The second question is more practical: Do spatial tests remain valid predictors <u>after</u> the ASVAB subtests have been entered into the equations. That is, would these tests lead to an overall improvement in the prediction of job performance if incorporated into Army-wide selection composites or would their predictive utility be limited to specific MOS and/or performance measures? These two research questions can be stated as follows:

1. How much variance in comprehensive performance measures can the Project A spatial tests alone account for, and how does this compare with ASVAB?

2. Which individual spatial tests remain significant predictors of comprehensive performance measures after the ASVAB subtests have been entered into the prediction equations?

Method

Subjects

Subjects were first-term enlisted personnel in nine entry-level MOS. These MOS had been selected for comprehensive examination in Project A because they were judged to be representative of the entire population of entry-level Army MOS. The sample consists of individuals who had entered the Army between 1 July 1983 and 30 July 1984 and was drawn from thirteen posts in the continental United States as well as the U.S. Army in Europe (USAREUR). The number of subjects from each MOS, as well as the total sample size, is shown in Table 1.

Table 1

Subjects for the Analyses

MOS	Enlisted Job	N	SQT
11B	Infantry	491	444
13B	Cannon Crew	464	396
19E	Armor Crew	394	338
31C	Single Channel Radio Operator	289	248
63B	Light Wheel Vehicle Mechanic	478	409
64C (now 88M)	Motor Transport Operator	507	427
71L	Administrative Specialist	427	361
91A	Medical Specialist	392	0
95B	Military Police	597	545
TOTAL		4,039	3,168

Note. Sample sizes shown in SQT column were those available for the analysis of the Skill Qualification Test criteria, as described later in this paper.

Predictors

Predictors were the nine ASVAB subtests and the six Project A paper-and-pencil tests of spatial ability. Detailed information concerning the nature and development of the Project

A predictors can be found in Peterson (1987). Table 2 presents a list of these predictors.

Table 2

Predictors Used in the Analyses

ASVAB Subtests:	Spatial Ability Tests:
Arithmetic Reasoning	Assembling Objects
Auto/Shop Information	Figural Reasoning
Coding Speed	Map
Electronics Information	Maze
General Science	Object Rotation
Math Knowledge	Orientation
Mechanical Comprehension	
Number Operations	
Verbal (Paragraph Comprehension + Word Knowledge)	

Criterion Measures

All criteria included in these analyses are comprehensive, "can-do" measures of proficiency in performing job duties, as described below.

Total Written Score on School and Job Knowledge Tests. School and job knowledge tests are multiple-choice measures of soldiers' technical knowledge pertinent to the various tasks performed in each MOS. Each school knowledge, or end-of-training test consists of 130-210 items, depending upon MOS (Davis, Davis, Joyner, & de Vera, 1987). The job knowledge tests were designed to measure knowledge of between 25 and 31 critical tasks and consist of 150-200 items, once again depending upon the particular MOS (Campbell, in preparation). Items on the job knowledge tests were selected to sample content broadly and to discriminate maximally among examinees. The total score is a unit-weighted composite of standard scores on the school and job knowledge tests.

General Soldiering Proficiency. General Soldiering and Core Technical Proficiency are the two performance constructs which the Project A staff created to account for the variance in "can-do", or maximal, performance. General Soldiering Proficiency is a composite score on a variety of tasks common to many MOS (e.g., determining grid coordinates on maps, recognizing friendly vs. threat aircraft), as measured by written test items and hands-on tasks. (MOS 11B was not scored on this measure.)

Core (i.e., MOS-specific) Technical Proficiency. This is defined as soldiers' ability to perform the tasks that are at the "core" of each MOS (i.e., those that define the MOS), each score being an aggregate of scores on written test items and hands-on tasks.

Skill Qualification Test Score (SQT). These are paper-and-pencil tests of MOS-specific technical knowledge developed by the U.S. Army Training and Doctrine Command for periodic testing of soldiers in their MOS. This was the only criterion analyzed which was not developed in Project A. (MOS 91A was not scored on this measure.)

It should be stressed that these measures are not mutually exclusive. That is, written test scores were included in the computation of the two composite measures, General Soldiering and Core Technical Proficiency, as well as the Total Written Score.

Procedure

Collection of Project A predictor and criterion data was part of the 1985 concurrent validation which occurred between 10 June and 13 November of that year. Scores on the ASVAB subtests and the Skill Qualification Test were obtained from archival data sources.

To answer the questions raised in the introduction, a series of backward stepwise multiple regression analyses were performed separately for each MOS. This procedure first enters all specified predictors into the equation as a block, then removes nonsignificant predictors one by one, based upon their individual contributions to the overall R^2. This backward elimination procedure continues until all variables in the equation are individually significant at a certain probability level (the present analyses used .05), and thus make a unique contribution to the prediction. Employing an SPSS Regression program, the analyses proceeded in two stages:

(1) SPATIAL ---> (SPATIAL)

The six spatial tests were entered as a block into each equation and the backward procedure determined which were individually significant.

(2) (SPATIAL)+ASVAB ---> (SPATIAL+ASVAB)

Significant spatial tests were retained in a second stage of analysis, meaning that they were no longer subject to removal. The ASVAB subtests were entered as a block and retained only if they added individually significant variance to the R^2 obtained for the spatial tests alone.

Busciglio (1990) used the following Stage (1) procedure:

(1) ASVAB ---> (ASVAB)

The nine ASVAB subtests were entered as a block into each equation and the backward procedure determined which were individually significant.

As can be seen, a comparison of the R^2s obtained from the present analyses to those of Busciglio (1990) will show the relative predictive power of spatial and ASVAB tests for predicting comprehensive performance measures. The Stage (2) results of the present analyses will show which spatial predictors <u>remain</u> significant after the ASVAB subtests are entered, thus indicating which can be expected to increment, not duplicate ASVAB validity.

The Army's selection and classification decisions are based on the ASVAB subtests. Because of this, restriction of range in the ASVAB scores (as well as any implicit range restriction in the measures which are correlated with them) would probably lead to underestimates of the actual criterion variance which would have been accounted for if ASVAB scores had not been used as a selection screen. To correct for this problem, matrices of estimated population (i.e., unrestricted) covariances among predictors and criteria were created, using the Lawley formula presented in Lord and Novick (1968; pp. 184-188), and used as input in these analyses. The R^2s reported below are in terms of this unrestricted population (the 1980 youth population, composed of individuals between the ages of 18 and 23) and have also been adjusted for shrinkage, using the formula given by Wherry (1940).

Results

First Research Topic

Table 3 shows proportions of criterion variance explained (R^2s) for groups of significant predictors at Stage (1) of the present analyses and those by Busciglio (1990). As can be seen, in many comparisons the spatial tests account for approximately as much criterion variance as do the ASVAB tests. The exceptions are shown in Table 4, which lists the criterion measures and MOS where the difference in R^2s for the two groups of tests equals or exceeds .05 (an arbitrary cutoff). As the table shows, the ASVAB tests were superior to the spatial measures in all instances, and were clearly better predictors of the two written criteria - Total Written Score on School and Job Knowledge Tests and the Skill Qualification Test. The superiority of the ASVAB tests for predicting General Soldiering and Core Technical Proficiency, which are composites of written and hands-on measures, was limited to a smaller number of MOS.

Table 3

Proportions of Variance Explained (R^2s) for Groups of Significant Spatial (SP) and ASVAB (ASV)[a] Tests at Stage 1 of Analyses

Criteria:	Total Written		General Soldiering		Core Technical		Skill Qualification	
Predictors:	(SP)	(ASV)	(SP)	(ASV)	(SP)	(ASV)	(SP)	(ASV)
<u>MOS</u>								
11B	.59	.59	--	--	.48	.48	.38	.44
13B	.40	.39	.30	.30	.16	.15	.19	.19
19E	.57	.58	.47	.44	.32	.35	.39	.40
31C	.50	.60	.45	.49	.38	.54	.44	.55
63B	.50	.63	.31	.28	.32	.48	.40	.55
64C	.49	.55	.46	.49	.28	.32	.38	.50
71L	.55	.54	.44	.41	.42	.44	.43	.58
91A	.54	.67	.49	.54	.45	.58	--	--
95B	.58	.62	.56	.59	.39	.43	.45	.60
MEDIAN	.54	.59	.46	.47	.38	.44	.40	.53

[a] data from Busciglio (1990).

Table 4

Largest Differences in R^2s for Groups of Significant Spatial and ASVAB Tests at Stage 1 of Analyses

Criteria:	Total Written		General Soldiering		Core Technical		Skill Qualification	
Predictors:	(SP)	(ASV)	(SP)	(ASV)	(SP)	(ASV)	(SP)	(ASV)
<u>MOS</u>								
11B			--	--			.38	.44
13B								
19E								
31C	.50	.60			.38	.54	.44	.55
63B	.50	.63					.40	.55
64C	.49	.55					.38	.50
71L							.43	.58
91A	.54	.67	.49	.54	.45	.58	--	--
95B							.45	.60
MEDIAN	.54	.59			.38	.44	.40	.53

<u>Note</u>. Differences shown are greater than or equal to .05, an arbitrary cutoff.

Second Research Topic

Table 5 lists the individual spatial tests which attained significance in Stage 1 and notes with an asterisk those which remained significant after the ASVAB tests were entered (and nonsignificant ones removed) in Stage 2. A summary of these results below shows the number of times (out of a maximum of 34) that each spatial test was significant in Stage 1 and remained significant in Stage 2:

Spatial Test	Stage 1	Stage 2
Assembling Objects	23	21
Figural Reasoning	27	20
Map	34	18
Maze	6	0
Object Rotation	5	2
Orientation	16	1

As can be seen, Assembling Objects, Figural Reasoning, and Map were especially strong predictors of the criterion measures across MOS. Assembling Objects and Figural Reasoning remained significant in the vast majority of cases. The Map test, while the strongest of all predictors in Stage 1, remained significant a smaller percentage of the time. The other three spatial tests were much weaker overall and remained significant in only a very few cases.

Discussion

Findings pertinent to the first research topic indicated that the spatial tests measure abilities which are important across a wide variety of MOS and criteria, being comparable to ASVAB as predictors of the two composite criteria. Although the spatial tests were somewhat inferior as predictors of the two written performance measures, it should be remembered that the spatial tests, while in a written format, probably do not tap as much verbal or "scholastic" ability as do the ASVAB tests.

An ancillary finding was that the superiority of ASVAB over spatial tests for predicting comprehensive performance may be, to some extent, a function of MOS. For example, ASVAB was superior to spatial tests for predicting three of the four criterion measures in two enlisted jobs, 31C and 91A (see Table 4). Further research is needed to delineate the job characteristics that determine the extent to which performance is differentially predictable from ASVAB vs. spatial measures.

Referring to the second research topic, three of the spatial tests - Assembling Objects, Figural Reasoning, and Map - were very strong predictors across MOS and criteria and tended to remain significant after the ASVAB tests were entered and deleted

Table 5

Significant Spatial Predictors at Stage 1 of Analyses

<u>Total Written Score</u>:						
11B		Fig Reas*	Map*			Orient
13B	Assem Objs*		Map*		Obj Rota	
19E	Assem Objs*	Fig Reas*	Map*			
31C		Fig Reas*	Map	Maze		
63B	Assem Objs*	Fig Reas*	Map*			Orient
64C	Assem Objs*	Fig Reas*	Map			
71L	Assem Objs*	Fig Reas*	Map*			
91A	Assem Objs*	Fig Reas*	Map			
95B	Assem Objs*	Fig Reas*	Map*			Orient
<u>General Soldiering Proficiency</u>:						
13B	Assem Objs*		Map*	Maze		
19E	Assem Objs*	Fig Reas*	Map*			
31C		Fig Reas*	Map	Maze		
63B	Assem Objs*		Map*			Orient
64C	Assem Objs*	Fig Reas*	Map		Obj Rota*	Orient*
71L	Assem Objs*	Fig Reas	Map		Obj Rota	Orient
91A	Assem Objs*	Fig Reas*	Map			Orient
95B	Assem Objs*	Fig Reas*	Map*			Orient
<u>Core Technical Proficiency</u>:						
11B		Fig Reas*	Map*			Orient
13B	Assem Objs*		Map*			
19E	Assem Objs*		Map*		Obj Rota*	
31C		Fig Reas	Map	Maze		
63B	Assem Objs*		Map			Orient
64C	Assem Objs*	Fig Reas*	Map			
71L	Assem Objs*	Fig Reas*	Map*		Obj Rota	
91A	Assem Objs	Fig Reas*	Map			
95B		Fig Reas*	Map*			Orient
<u>Skill Qualification Test Score</u>:						
11B		Fig Reas*	Map*			Orient
13B	Assem Objs*		Map*			
19E	Assem Objs*	Fig Reas	Map*			
31C		Fig Reas	Map	Maze		Orient
63B	Assem Objs	Fig Reas	Map			Orient
64C		Fig Reas	Map	Maze		Orient
71L		Fig Reas*	Map			
95B		Fig Reas	Map			Orient

* Remained significant after entry of ASVAB tests in Stage 2.
Assem Objs - Assembling Objects   Obj Rota - Object Rotation
Fig Reas - Figural Reasoning   Orient - Orientationin

in Stage 2. This supports the notion that these tests would lead to an overall improvement in the prediction of job performance if incorporated into the Army's selection composites. In contrast, the usefulness of the Maze, Object Rotation, and Orientation tests seems to be much more limited to specific MOS and/or performance measure. Thus, their greatest utility would probably be in MOS-specific classification composites.

To interpret these results properly, it is important to note a number of methodological considerations. As stated earlier, individual written tests contributed variance to the total written score on school and job knowledge tests as well as to General Soldiering and Core Technical Proficiency. Thus, results across these criteria are not independent. Also, possible differences in motivation due to differences in testing situations (i.e., ASVAB scores used for selection as opposed to Project A scores used "for research purposes only") may have impacted the results. That is, individuals may have responded more carefully, exerted more effort, etc., on the ASVAB subtests, thus making them more valid measures of abilities than the Project A tests. The final methodological concern has to do with the statistical analyses used. Stepwise regression procedures, while useful for empirically exploring alternative models, are especially susceptible to sampling error (cf. Cohen & Cohen, 1983). The samples used in the analyses were generally of sufficient size to make the degree of shrinkage in each individual equation relatively low, but the large number of equations computed here increases the probabilities that some ASVAB and Project A predictors are significant due to Type I errors. However, it should be noted that each predictor could have been significant a maximum of 34 times, across all criteria and MOS. At an alpha level of .05, a predictor may be expected to reach significance, by sampling error alone, between one and two times. As Table 5 clearly indicates, most of the Project A tests were significant far more often than this. Nevertheless, the lack of opportunities at this point for cross-validation renders the results reported in this paper suggestive only.

References

Armed Services Vocational Aptitude Battery (ASVAB) (1984). Information Pamphlet. DOD 1304.12Z.

Busciglio, H.H. (1990). The incremental validity of spatial and perceptual-psychomotor tests relative to the Armed Services Vocational Aptitude Battery (ARI Technical Report 883). Alexandria, VA: U.S. Army Research Institute. (AD A220 903)

Campbell, C.H. (in preparation). Developing basic criterion scores for hands-on tests, job knowledge tests, and task rating scales (Draft of ARI Technical Report).

Campbell, J.P. (Ed.). (1988). Improving the selection, classification, and utilization of Army enlisted personnel: Annual report, 1986 fiscal year (ARI Technical Report 792). Alexandria, VA: U.S. Army Research Institute. (AD A198 856)

Cohen, J., & Cohen, P. (1983). Applied multiple regression/ correlation analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Davis, R.H., Davis, G.A., Joyner, J.N., & de Vera, M.V. (1987). Development and field test of job relevant knowledge tests for selected MOS (ARI Technical Report 757). Alexandria, VA: U.S. Army Research Institute. (AD A192 211)

Lord, P., & Novick, M. (1968). Statistical theory of mental test scores. Reading, MA: Addison-Wesley Publishing Co.

Peterson, N.G. (Ed.). (1987). Development and field test of the trial battery for Project A (ARI Technical Report 739). Alexandria, VA: U.S. Army Research Institute. (AD A184 575)

Wherry, R.J. (1940). Appendix A. In W.H. Stead and C.P. Shartle (Eds.), Occupational counseling techniques. New York: American Book Company.